

09/06/00

Please type a plus sign (+) inside this box →

Approved for use through 09/30/2000 OMB 0651-0032

Patent and Trademark Office U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

**UTILITY
PATENT APPLICATION
TRANSMITTAL**

(Only for new nonprovisional applications under 37 C.F.R. § 1.53(b))

Attorney Docket No. 367.39020X00

First Inventor or Application Identifier Mika KIVIMAKI

Title See 1 in Addendum

Express Mail Label No. _____

P-10

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

APPLICATION ELEMENTS

See MPEP chapter 600 concerning utility patent application contents.

1. Fee Transmittal Form (e.g., PTO/SB/17)
(Submit an original and a duplicate for fee processing)
2. Specification [Total Pages 23]
(preferred arrangement set forth below)
 - Descriptive title of the Invention
 - Cross References to Related Applications
 - Statement Regarding Fed sponsored R & D
 - Reference to Microfiche Appendix
 - Background of the Invention
 - Brief Summary of the Invention
 - Brief Description of the Drawings (if filed)
 - Detailed Description
 - Claim(s)
 - Abstract of the Disclosure
3. Drawing(s) (35 U.S.C. 113) [Total Sheets 7]
4. Oath or Declaration [Total Pages]
- a. Newly executed (original or copy)
- b. Copy from a prior application (37 C.F.R. § 1.63(d))
(continuation/divisional with Box 15 completed)
 - i. DELETION OF INVENTOR(S)
Signed statement attached deleting inventor(s) named in the prior application, see 37 C.F.R. §§ 1.63(d)(2) and 1.33(b).

*** NOTE FOR ITEMS 1 & 13 IN ORDER TO BE ENTITLED TO PAY SMALL ENTITY FEES, A SMALL ENTITY STATEMENT IS REQUIRED (37 C.F.R. § 1.27), EXCEPT IF ONE FILED IN A PRIOR APPLICATION IS RELIED UPON (37 C.F.R. § 1.28).**

16. If a CONTINUING APPLICATION, check appropriate box, and supply the requisite information below and in a preliminary amendment:

Continuation Divisional Continuation-in-part (CIP) of prior application No. / _____

Prior application information: Examiner _____

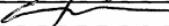
Group / Art Unit: _____

For CONTINUATION or DIVISIONAL APPS only: The entire disclosure of the prior application, from which an oath or declaration is supplied under Box 4b, is considered a part of the disclosure of the accompanying continuation or divisional application and is hereby incorporated by reference. The incorporation can only be relied upon when a portion has been inadvertently omitted from the submitted application parts.

17. CORRESPONDENCE ADDRESS

<input checked="" type="checkbox"/> Customer Number or Bar Code Label	020457	<input type="checkbox"/> Correspondence address below
(Insert Customer No. or Attach bar code label here)		

Name			
Address			
City	State	Zip Code	Fax
Country			

Name (Print/Type)	Carl I. Brudidge	Registration No. (Attorney/Agent)	29,621
Signature			
	Date 9-06-00		

Burden Hour Statement: This form is estimated to take 0.2 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Box Patent Application, Washington, DC 20231.

Attachment to PTO/SB/05 (4/98) Utility Patent Application
Transmittal

1. USER INTERFACE FOR TEXTTO SPEECH CONVERSION
 2. - FIGS. 1-7
 3. - CREDIT CARD PAYMENT FORM

卷之三

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

FEE TRANSMITTAL for FY 2000

Patent fees are subject to annual revision.
Small Entity payments must be supported by a small entity statement,
otherwise large entity fees must be paid. See Forms PTO/SB/09-12
See 37 C.F.R. §§ 1.27 and 1.28.

TOTAL AMOUNT OF PAYMENT (\$1,422.00)

Complete if Known

Application Number	
Filing Date	September 6, 2000
First Named Inventor	Mika KIVIMAKI
Examiner Name	
Group / Art Unit	
Attorney Docket No.	367.39020X00

METHOD OF PAYMENT (check one)

- The Commissioner is hereby authorized to charge indicated fees and credit any overpayments to:

Deposit Account Number **01-2135**

Deposit Account Name **Antonelli, Terry, Stout & Kraus, LLP**

Charge Any Additional Fee Required
Under 37 CFR §§ 1.16 and 1.17

Payment Enclosed:

Check Money Other

FEES CALCULATION**1. BASIC FILING FEE**

Large Entity	Small Entity	Fee Code (\$)	Fee Code (\$)	Fee Description	Fee Paid
101	690	201	345	Utility filing fee	690.00
106	310	206	155	Design filing fee	
107	480	207	240	Plant filing fee	
108	690	208	345	Reissue filing fee	
114	150	214	75	Provisional filing fee	
SUBTOTAL (1)				(\$ 690.00)	

2. EXTRA CLAIM FEES

Extra Claims	Fee from below	Fee Paid
Total Claims 39	-20** 19	X 18 + 342
Independent 8	-3** 5	X 78 + 190
Claims		
Multiple Dependent		0

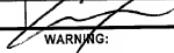
*or number previously paid, if greater; For Reissues, see below

Large Entity	Small Entity	Fee Code (\$)	Fee Code (\$)	Fee Description
103	18	203	9	Claims in excess of 20
102	78	202	39	Independent claims in excess of 3
104	280	204	130	Multiple dependent claim, if not paid
109	78	209	39	** Reissue independent claims over original patent
110	18	210	9	** Reissue claims in excess of 20 and over original patent
SUBTOTAL (2)				(\$ 732.00)

FEES CALCULATION (continued)**3. ADDITIONAL FEES**

Large Entity	Small Entity	Fee Code (\$)	Fee Code (\$)	Fee Description	Fee Paid
105	130	205	65	Surcharge - late filing fee or oath	0.00
127	50	227	25	Surcharge - late provisional filing fee or cover sheet	0.00
139	130	139	130	Non-English specification	0.00
147	2,520	147	2,520	For filing a request for reexamination	0.00
112	920*	112	920*	Requesting publication of SIR prior to Examiner action	0.00
113	1,840*	113	1,840*	Requesting publication of SIR after Examiner action	0.00
115	110	215	55	Extension for reply within first month	0.00
116	380	216	190	Extension for reply within second month	0.00
117	870	217	435	Extension for reply within third month	0.00
118	1,380	218	680	Extension for reply within fourth month	0.00
128	1,850	228	925	Extension for reply within fifth month	0.00
119	300	219	150	Notice of Appeal	0.00
120	300	220	150	Filing a brief in support of an appeal	0.00
121	280	221	130	Request for oral hearing	0.00
138	1,510	138	1,510	Petition to institute a public use proceeding	0.00
140	110	240	55	Petition to revive - unavoidable	0.00
141	1,210	241	605	Petition to revive - unintentional	0.00
142	1,210	242	605	Utility issue fee (or reissue)	0.00
143	430	243	215	Design issue fee	0.00
144	580	244	290	Plant issue fee	0.00
122	130	122	130	Petitions to the Commissioner	0.00
123	50	123	50	Petitions related to provisional applications	0.00
126	240	126	240	Submission of Information Disclosure Stmt	0.00
581	40	581	40	Recording each patent assignment per property (times number of properties)	0.00
146	690	246	345	Filing a submission after final rejection (37 CFR § 1.129(a))	0.00
149	690	249	345	For each additional invention to be examined (37 CFR § 1.129(b))	0.00
Other fee (specify) _____					0.00
Other fee (specify) _____					0.00
Reduced by Basic Filing Fee Paid					SUBTOTAL (3) (\$ 0.00)

SUBMITTED BY

Name (Print/Type)	Carl L. Brundidge	Registration No. (Attorney/Agent)	29,621	Telephone	703-312-6600
Signature		Date	9-06-00		

WARNING:

Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

Burden Hour Statement: This form is estimated to take 0.2 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Washington, DC 20231.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants: Mika KIVIMÄKI
Serial No.: Not yet assigned
Filed: September 6, 2000
For: USER INTERFACE FOR TEXT TO SPEECH CONVERSION
Group: Not yet assigned
Examiner: Not yet assigned

PRELIMINARY AMENDMENT

Assistant Commissioner for Patents
Washington, D.C. 20231

September 6, 2000

Sir:

Prior to examination, please amend the above-identified application as follows.

IN THE SPECIFICATION

Please amend the specification as follows:

Page 16, line 19, delete "What is claimed is:".

IN THE CLAIMS

Page 17, line 1, insert --What is claimed is:--.

Please cancel claim 29 without prejudice or disclaimer of the matter therein

Please amend claims 3, 4, 6-8, 9-18, 20-25, 28 and 33 as follows:

3. (Amended) A device as claimed in claim 1, [or 2] wherein the control means varies the highlighting with the conversion of text to audio output.

4. (Amended) A device as claimed in [any preceding] claim 1, wherein the control means extends the highlighting through the displayed text with the conversion of text to audio output.

6. (Amended) A device as claimed in claim 4 [or 5], wherein the extending of highlighting through a text portion lags the conversion of the text portion to audio output by the delay.

7. (Amended) A device as claimed in [any one of claims] claim 4 [to 6], wherein the highlighting extends discontinuously by portions of text corresponding to a word or words.

8. (Amended) A device as claimed in [any preceding] claim 1, wherein the control means highlights a text portion for a limited duration.

10. (Amended) A device as claimed in [any preceding] claim 1, wherein the dimension of the highlighted portion is variable.

11. (Amended) A device as claimed in [any preceding] claim 1, wherein the dimension of the highlighted portion is a constant number of words long.

12. (Amended) A device as claimed in claim 1, wherein the display displays the full sentence of text being converted.

13. (Amended) A device as claimed in [any preceding] claim 1, wherein the display displays the previous predetermined plurality of words that have been converted.

14. (Amended) A device as claimed in [any preceding] claim 1, further comprising:

a dictionary, wherein the control means highlights words not in the dictionary.

15. (Amended) A device as claimed in claim 14, wherein the control means provides different highlighting for the words not in the dictionary.

16. (Amended) A device as claimed in [any preceding] claim 1, wherein the control means is arranged to identify proper nouns in the text, and highlight them.

17. A device as claimed in claim 16, wherein the control means provides different highlighting for the proper nouns in the text.

18. (Amended) A device as claimed in [any preceding] claim 1, wherein the speech synthesizer means provides signals to the control means to effect the highlighting of specific words.

20. (Amended) A device as claimed in [any preceding] claim 1, wherein highlighting may comprise anyone of:
 contrast variation of the text and/or display background;
 [colour] color variation of the text and/or display background;
 reformatting of the text including underscoring, bold font, italic font, [capitalisation] capitalization, changing font size or type; and
 enclosing text in geometric shapes.

21. (Amended) A device as claimed in [any preceding] claim 1, wherein the delay corresponds to a number of converted words or syllables.

22. (Amended) A device as claimed in [any preceding] claim 1, wherein the delay corresponds to a fixed time.

23. (Amended) A device as claimed in [any preceding] claim 1, wherein the delay is greater than 0.1 seconds.

24. (Amended) A device as claimed in [any preceding] claim 1, wherein the delay is less than 3 seconds.

25. (Amended) A device as claimed in [any preceding] claim 1, wherein the electronic device is a document reader or a hand-held communications device.

28. (Amended) A car [having] comprising:

a hand-held radio communications device which comprises:

a user interface which includes:

a display for displaying text, and

speech synthesizer means including a loudspeaker,

arranged to convert an input, dependent upon a text, to an audio output representative of a person reading the text; and

control means for controlling the display and for providing an input to the speech synthesizer means, arranged

to control the display of a text, to provide an input

corresponding to the displayed text to the speech synthesizer,

and to highlight a portion or portions of the displayed text,

wherein the highlighting of a text portion is delayed with

respect to the audio output corresponding to the text portion

[device as claimed in claim 26 or] , or

[document reader as claimed in claim 27,]

a document reader which comprises:

a user interface having:

a display for displaying text, and

speech synthesizer means including a loudspeaker,

arranged to convert an input, dependent upon a text, to an audio output representative of a person reading the text; and

control means for controlling the display and for providing an input to the speech synthesizer means, arranged

to control the display of a text, to provide an input

corresponding to the displayed text to the speech synthesizer,

and to highlight a portion or portions of the displayed text,

wherein the highlighting of a text portion is delayed with

respect to the audio output corresponding to the text portion;[,]
wherein the display is mounted for use by the
driver.

33. (Amended) A device as claimed in claim 31, [or a
method as claimed in claim 32 where] wherein the delay is
greater than 0.1 seconds.

Please add new claims 34-39 as follows:

-- 34. A method as claimed in claim 32, wherein the delay
is greater than 0.1 seconds.

35. A device as claimed in claim 2, wherein the control
means varies the highlighting with the conversion of text to
audio output.

36. A device as claimed in claim 5, wherein the
extending of highlighting through a text portion lags the
conversion of the text portion to audio output by the delay.

37. A device as claimed in claim 5, wherein the
highlighting extends discontinuously by portions of text
corresponding to a word or words.

38. A device as claimed in claim 6, wherein the
highlighting extends discontinuously by portions of text
corresponding to a word or words.

39. A device as claimed in claim 36, wherein the highlighting extends discontinuously by portions of text corresponding to a word or words.--

IN THE ABSTRACT

Line 4, delete "comprising" insert --which includes--.

REMARKS

Entry of the above amendments prior to examination is respectfully requested.

Please charge any shortage in fees due in connection with the filing of this paper, or credit any overpayment of fees, to the deposit account of Antonelli, Terry, Stout & Kraus, LLP, Deposit Account No. 01-2135 (367.39020X00).

Respectfully submitted,

ANTONELLI, TERRY, STOUT & KRAUS, LLP



Carl I. Brundidge
Registration No. 29,621

CIB/jdc
(703) 312-6600

User interface for text to speech conversion

Background of the Invention

5

The present invention relates to a user interface for providing a text output on a display and an audio output corresponding to the displayed text.

10

The synthesis of human speech using electronic devices is a well developed and published technology and various commercial products are available. Typically speech synthesis programs convert written input to spoken output by automatically generating synthetic speech and speech synthesis is therefore often referred to a "Text-to-Speech" conversion (TTS).

20

There are several problems in speech synthesis which, as yet, have not been satisfactorily resolved and which impact on the comprehension of the synthetic speech by a user. One problem is that certain words and syllables are more difficult to pronounce than others. Another problem is that the pronunciation rules are language dependent and a French word in an English text will often be mispronounced by the speech synthesiser.

The above mentioned problems may be exacerbated in mobile electronic devices such as mobile telephones or pagers which may have limited processing resources

25

It would be desirable to improve the level of comprehension a user has of the speech output from such speech synthesiser systems.

Summary of the Invention

30

According to one aspect of the present invention there is provided an electronic device comprising a user interface having: a display for displaying text and speech synthesiser means including a loudspeaker, arranged to convert an input, dependent upon a text, to an audio output representative of a person
5 reading the text; and control means, for controlling the display and for providing an input to the speech synthesiser means, arranged to control the display of a text, to provide an input corresponding to the displayed text to the speech synthesiser, and to highlight a portion or portions of the displayed text, wherein the highlighting of a text portion is delayed with respect to the audio output
10 corresponding to the text portion.

Embodiments of the present invention therefore provide a display on which a user can check a word which has been mispronounced by the speech synthesiser. Moreover, the highlighting draws the user's eye to the relevant portion of displayed text at the correct time. The delay allows time for a user to register that a synthesised word was not understood and to look to the display.

The highlighting of text is particularly important when a user is performing another task while listening to the synthesised speech. The user is directed to the relevant portion of text with the minimum of distraction. The user does not need to scan unnecessary text to find the relevant portion of text. The user therefore expends less time and effort in comprehending the synthesised speech.

25 The control means may synchronise the highlighting with the conversion of text to audio output and/or may vary the highlighting with the conversion of text to audio output and/or may extend the highlighting selectively or unselectively through the displayed text with the conversion of text to audio output.

The extending of highlighting through a text portion may lag the conversion of the text portion to audio output by the delay. The highlighting may extend discontinuously by portions of text corresponding to a word or words. Highlighting of a text portion may be for a limited duration.

5

The highlighting may therefore isolate a portion or portions of text from a body of text, the isolation of text lagging the conversion of the text to audio output by the delay and the isolation being maintained for the limited duration.

10 The dimension of the highlighted portion may be variable or fixed. The dimensions of the highlighted portion may be a constant number of words, characters or syllables long.

15 The display may display the full sentence of text being converted. The display may display the previous predetermined plurality of words that have been converted.

20 The device may additionally comprise a dictionary, wherein the control means highlights words that are not in the dictionary. This highlighting may be different to the previously referred to highlighting.

25

The control means may be arranged to identify proper nouns in the text, and highlight them. This highlighting may be different to the previously referred to highlighting.

The synthesiser means may provide signals to the control means to effect the highlighting of specific words. This highlighting may be different to the previously referred to highlighting.

The highlighting may be any way of visually differentiating the highlighted text from the not highlighted text. Contrast variation of the text and/or display background, colour variation of the text and/or display background; reformatting of the text including underscoring, using bold font, italic font or capitalisation and
5 changing font size or type; and enclosing the text in geometric shapes may in any combination be used to highlight text.

The delay may correspond to a fixed number of converted words or syllables.

The delay may correspond to a fixed time. The delay is preferably greater than

10 0.1 seconds. The delay may be less than 3 seconds.

15 The electronic device may be portable and is preferably hand-held.

The electronic device may be a radio communications device such as mobile radio telephone. The electronic device may be a document reader such as a Web browser, a Text-TV, a fax, a document browser for reading books, emails or other documents.

20 According to another aspect of the present invention there is provided a method

for displaying text and providing speech synthesis of the text comprising the steps of: displaying the text; converting a text portion to audio output; determining that the text portion should be highlighted; delaying; and then highlighting the text portion. The display of a text may be already displayed when the speech corresponding to the text is synthesised, it may be simultaneous with
25 the synthesis of that speech or it may be delayed with respect to the synthesis of that speech.

According to another, different, aspect of the present invention there is provided

an electronic device comprising: a user interface having: a display for displaying

30 text and speech synthesiser means including a loudspeaker, arranged to convert

an input, dependent upon a text, to an audio output representative of a person reading the text; and control means, for controlling the display and for providing an input to the speech synthesiser means, arranged to control the display of a text, to provide an input corresponding to the displayed text to the speech synthesiser, and to delay the display of a text portion with respect to the audio output corresponding to the text portion. This delay is preferably greater than 0.1 seconds. There is also provided a method for displaying text and providing speech synthesis of the text comprising the steps of: converting a text portion to audio output; delaying; and then displaying the text.

10

Brief Description of the Drawings

For a better understanding of the present invention and how it may be brought into effect reference will now be made by way of example only to the accompanying Figures in which:

Figure 1 illustrates an electronic device with a user interface having a loudspeaker and display

20 Figure 2 is a schematic illustration of the components of the electronic device illustrated in Figure 1

Figure 3 illustrates the unselective highlighting of portions of a body of text;

25 Figure 4 illustrates the selective highlighting of portions of a body of text which are not in a dictionary;

Figure 5 illustrates the selective highlighting of portions of a body of text which are proper nouns;

30

Figure 6 illustrates the selective highlighting of portions of a body of text which are identified as difficult to convert accurately by the text-to-speech synthesiser; and

- 5 Figure 7 illustrates an example database for controlling delays.

Detailed Description of the Drawings

Figure 1 illustrates an electronic device 2. The electronic device communicates 10 to a user through its user interface which includes a display 4 and a loudspeaker 6. The device broadcasts synthesised speech to a user through the speaker 6 and at the same time displays text, from which the speech is derived, on the display. A user who does not understand the broadcast speech can improve their comprehension by reading the display. Portions of the display are highlightable 15 to draw the users attention to those portions. Highlighting of text only occurs after the loudspeaker has broadcast the synthesised speech corresponding to the text. The electronic device may be any device which requires an audio-visual interface. It may be a personal digital assistant (PDA), a radio communications device such as mobile radio telephone e.g. a car-phone or hand-held phone, a 20 computer system, a document reader such as a Web browser, a Text-TV, a fax, a document browser for reading books, emails or other documents or the like.

Although the display 4 and loudspeaker 6 in Figure 1 are shown as being integrated in a single unit they may be separate.

25 Figure 2 is a schematic illustration of the electronic device 2. The device 2 in addition to having the display 4 and the loudspeaker 6 has a processor 12 for driving the display and the loudspeaker and for accessing a memory 10. The memory 10 stores text data supplied via an input 30. The processor 12 is 30 illustrated as two functional blocks- a controller 14 and a text-to-speech engine

16. The controller 14 and text-to-speech engine 16 may be implemented as software running on the processor 12.

The text-to-speech engine 16 drives the loudspeaker 6. It receives a text input 18 from the controller and converts the text input to a synthetic speech output 22 which is transduced by the speaker 6 to sound waves. The speech output may be one word at a time, one phrase at a time or one sentence at a time.

10 The controller 14 reads the memory 20 and controls the display 4 and the text-to-speech engine 16. The controller having read text data from the memory provides it as an input 18 to the text-to-speech engine 16 and controls the display to display the text data using control signals 16. The displayed text corresponds to the text converted by the text-to-speech engine 16.

15 The memory 10 stores text data which is read by the controller 14. The controller 14 uses the text data to produce the input 18 to the text-to-speech engine 16 and the control signals 16 to the display. Text data is stored in the memory 10 by the input device 30. The input device in this example includes a radio transceiver for receiving text data in the form of SMS messages or emails.

20 The controller also controls the display to highlight text using the control signals 16. Highlighting of text includes anything which visually differentiates a portion of text from the remaining body of text on the display 4. Contrast variation of the text and/or display background, colour variation of the text and/or display background, reformatting of the text including underscoring, using bold font, italic font or capitalisation and changing font size or type and enclosing the text in geometric shapes may in any combination be used to highlight text.

25 The highlighting of the text is synchronised to the conversion of the text to audio output.

As each portion of the text is converted to audio output the controller determines if that text portion should be highlighted, if it should be highlighted, it delays, then controls the highlighting of the text portion. A portion of text may be a character,
5 a syllable, a word, a phrase, a sentence or sentences.

The delay is preferably of the order of a second or seconds but may be varied. It is greater than 0.1 second and less than 3 seconds. In particular, the delay may be related to the speed of the text-to-speech processor, the delay increasing as
10 the speed decreases and visa-versa.

15 The controller 14 may be selective or unselective in the highlighting of text. Unselective highlighting is the highlighting of each portion of text.

20 The highlighted portion of text may have a fixed length or a variable length. A portion of text, once highlighted remains temporarily highlighted until un-highlighted by the controller 14. Typically, this will occur after a predetermined time. When the highlighting of text is unselective it may occur when the highlighting has extended beyond the portion to be un-highlighted by a predetermined amount. When the highlighting of text is selective it may occur when the controller determines if a text portion, ahead of the portion to be un-highlighted by a predetermined amount, should be highlighted.

25 In Figures 3, there is illustrated an example of unselective highlighting. In Figures 4, 5 and 6 are illustrated different examples of selective highlighting. In each of these examples, the text stored in the memory 10 and read by the controller 14 is "Thank-you for your email. I look forward to meeting tou [sic] this afternoon in London, a bientot." To the left of the Figure is illustrated the display 16, controlled by controller 14, on which the text is shown. To the right the audio out
30 put of the loudspeaker 6 is illustrated. The visual output to the left and the audio

output to the right evolve as the figure is descended. The time arrow to the left of the Figure marks the evolution.

The operation of the controller 14 will now be described in more detail. The controller accesses a text from the memory 10. It parses the text for selective highlighting purposes and for unselective highlighting and text to speech conversion. The parsing for selective highlighting purposes involves the identification of individual words by recognising space characters within the text.

The parsing for text to speech synthesis involves the separation of the text into separate text groups each of which includes one or several words. The word groups are sent one at a time to the TTS engine 16.

If the controller is performing selective highlighting it analyses each of the words identified in the text. It can separately identify and indicate whether the word is unusual or foreign, whether the word is a proper name or whether the TTS engine is likely to mispronounce the word. The controller has a dictionary for the language being used which may contain words of common usage. A comparison of the word with the dictionary and its absence from the dictionary indicate that the word is "unexpected" and it can be selectively highlighted to indicate this.

The controller may identify proper names in the English and French languages for example by recognising capital font for letters which start a word but which do not follow a full stop (period). The controller can then identify the word for selective highlighting to indicate that the word is "proper". The controller may have a lexicon for the TTS which identifies words or letter groupings within words which may be incorrectly converted by the TTS engine or receive feed back from the TTS engine 16 which identifies a word group which is or is likely to be incorrectly converted by the TTS engine. The controller can then identify the word or the word or words in the group for selective highlighting to indicate that the word or group of words is "difficult".

The controller maintains a database to enable control of the delay in highlighting and/or text display. An example database is illustrated in Figure 7. For the purposes of the example the phrase "Thank-you for your email. I look forward to meeting you this afternoon in London, a bientot." has been parsed. The data base 30 has an entry for each of the 17 words in the phrase. Each entry has four fields. The first field 32 records the number of characters in a particular word entry. The second field 34 records which text group the word entry belongs to by storing a text group id which is different for each text group. The third field 36 records whether selective highlighting is required for the word and whether the word should be identified as "difficult", "unexpected" or "proper". The fourth field 38 records when the time or the expected time when the TTS engine 16 produces an audio output for the text group of which the entry is part.

After parsing the text input the entries for the first 32, second 34 and third 36 fields can be recorded for each entry. The input to the fourth field may be input predictively or responsively. For responsive input, when the controller sends a text group to the TTS engine it also sends the text group ID. The TTS engine stores the received text group IDs in the order in which they are received in a FIFO 16. When a text portion has been converted to speech it pops the FIFO 20 returns the text group ID identifying the text group which has just been converted as signal 18 to the controller. When the controller receives this ID it records the time from its internal clock in the fourth field against those entries which have the same text group ID in their second field as the returned text group ID. For predictive input, an approximation is used for the time taken on average to convert a word (or character) to speech from when it is sent to the TTS engine. This is a parameter of the TTS engine 16. From the database 30, the number of words (and characters) including punctuation in a text group can be calculated and an estimate of the time at which the text group will be converted to speech deduced. The expected time of conversion for each text group is deduced. For

text group the appropriate time is recorded the fourth field for the each entry which identifies that text group in its second field.

Referring to Figure 7 there are four discrete text groups having text group IDs 1,

- 5 2, 3 and 4. Group 1 includes the words/characters "Thank-you for your email." . Group 2 includes the words/characters " I look forward to meeting tou this afternoon". Group 3 includes the words/characters " in London.". Group 4 includes the words/characters " a bientot.". The text groups 1, 2, 3 and 4 are respectively converted to speech at times t1, t3, t5 and t7.

10

When the controller sends a text group to the TTS engine 16 it controls the display to display that text group. It may in certain circumstances be desirable to delay the display of the text so that it is not displayed in advance of the text being highlighted. This delay in displaying the text, if used, should be greater than 0.1 seconds but less than three seconds.

20

At a time $t_2=t_1 + T$, where T is a predetermined delay, the controller accesses the database and controls the display to highlight the characters comprising the text group 1 if unselective highlighting is activated. If selective highlighting of "difficult" words is activated, then at time t_2 , the controller accesses the database and controls the display to selectively highlight the word "email".

25

At a time $t_4=t_3 + T$, where T is a predetermined delay, the controller controls the display to highlight the characters comprising the text group 2 if unselective highlighting is activated. If selective highlighting of "unexpected" words is activated, then at time t_4 the controller accesses the database and controls the display to selectively highlight the word "tou".

- 30 At a time $t_6=t_5 + T$, where T is a predetermined delay, the controller accesses the database and controls the display to highlight the characters comprising the

text group 3 if unselective highlighting is activated. If selective highlighting of "proper" words is activated, then at time t6 the controller accesses the database and controls the display to selectively highlight the word "London".

- 5 At a time $t_8=t_7 + T$, where T is a predetermined delay the controller accesses the database and controls the display to highlight the characters comprising the text group 4 if unselective highlighting is activated. If selective highlighting of "unexpected" words is activated, then at time t8 the controller accesses the database and controls the display to selectively highlight the word "bientot".
- 10 If selective highlighting of "difficult " words is activated, then at time t8 the controller accesses the database and controls the display to selectively highlight the word "bientot".

The delay T can be adjusted and varied by the user to achieve their preferred delay. The delay T may be varied in dependence on the number of words and characters in a text group. In particular, the predetermined delay may be decreased as the text group size increases.

- 20 The removal of highlighting from a word or text group may be effected in a number of different ways. For example in unselective highlighting, highlighting may not be removed until the whole sentence in which the text group appears has been highlighted for at least a certain time. Alternatively highlighting may be removed from a text portion after it has been highlighted for a certain time. Another alternative is to have one or a certain number of contiguous text groups 25 highlighted at any one time where the highlighting of a new text group results in the removal of highlighting from an old text group and the progression of the highlighting window through the text. For selective highlighting the word may remain highlighted until the adjacent text has its highlighting removed or would have done so if it had been highlighted or it may be highlighted for a certain time.

- According to the first example (of unselective highlighting), illustrated in Figure 3, as each portion of the text is converted to audio output the controller determines by default that the text portion should be highlighted, it delays, then controls the highlighting of the text portion. The highlighting is temporary. The highlighting
5 extends at a leading edge through the displayed text one word or several words at a time and contracts at a trailing edge one word or several words at a time.

At time t0, the controller 14 places the text onto the display 4 and provides the text to the text-to-speech engine 16. At time t1, a portion of text, in this case the
10 sentence "Thank-you for your email." is converted to audio output. The displayed text is unaffected.

After a delay, at time t2, the controller 14 highlights the portion of text "Thank-you for your email." in the body of displayed text.

- 15 At time t3, another portion of text, in this case the phrase "I look forward to meeting tou this afternoon" is converted to audio output. The displayed text and highlighting is unaffected.
20 After a delay, at time t4, the controller 14 highlights the portion of text "I look forward to meeting tou this afternoon" in the body of displayed text. By the time t4, the controller has un-highlighted the sentence "Thank-you for your email."

25 At time t5, another portion of text, in this case the phrase "in London" is converted to audio output. The displayed text and highlighting is unaffected.

After a delay, at time t6, the controller 14 highlights the portion of text "in London" in the body of displayed text. At the time t6, the controller is still highlighted the phrase "I look forward to meeting tou this afternoon".

At time t7, another portion of text, in this case the phrase "a bientot" is converted to audio output, the displayed text and highlighting is unaffected.

After a delay, at time t8, the controller 14 highlights the portion of text "a bientot"

- 5 in the body of displayed text. By the time t8, the controller has un-highlighted the phrase "I look forward to meeting tou this afternoon".

According to a second example (of selective highlighting), illustrated in Figure 4, highlighting may be selectively applied to the displayed text. In this example, text
10 is highlighted if it is not in a dictionary.

The controller determines if a word in a text group should be highlighted, if it should be highlighted, it waits after the conversion of the text group to speech, then controls the highlighting of the text portion.

At time t0, the controller 14 places the text onto the display 4.A a portion of text, in this case the sentence "Thank-you for your email." is sent to the TTS engine for conversion to audio output and converted at time t1. The controller determines if that text portion should be highlighted in the displayed text. It checks each word in the sentence against the words stored in a dictionary. Each word is recognised and no highlighting is required. The displayed text is unaffected.

At time t3, a portion of text, in this case the phrase "I look forward to meeting tou this afternoon" is converted to audio output. The controller determines if that text portion should be highlighted in the displayed text. It checks each word in the sentence against the words stored in a dictionary. Each word except "tou" is recognised. "tou" is a typographical error and is not in the dictionary. The controller therefore determines that the text portion "tou" should be highlighted, it delays until time t4, then controls the highlighting of the text portion "tou".

At time t5 a portion of text, in this case the phrase "in London" is converted to audio output. The controller determines if that text portion should be highlighted in the displayed text. It checks each word in the sentence against the words stored in a dictionary. Each word is recognised and no highlighting is required.

At time t7, a portion of text, in this case the French phrase "a bientot," is converted to audio output. The controller determines if that text portion should be highlighted in the displayed text. It checks each word in the sentence against the words stored in a dictionary. "bientot" is not recognised in an English dictionary. The controller determines that the text portion "bientot" should be highlighted, it delays until time t8, then controls the highlighting of the text portion "bientot". By the time t8, the controller has un-highlighted "tou".

According to a third example, illustrated in Figure 5, highlighting may be selectively applied to the displayed text. In this example, text is highlighted if it is a proper name. As each portion of the text is converted to audio output the controller determines if that text portion should be highlighted, if it should be highlighted, it delays, then controls the highlighting of the text portion. The controller identifies words which begin with a capital letter but which are not following a full-stop (period). The word "London" is produced as audio output at time t5 and is highlighted, after a delay, at time t6. By time t8 the highlighting has been removed.

According to a fourth example, illustrated in Figure 6, highlighting may be selectively applied to the displayed text. In this example, text identified as problematic for the text-to-speech engine 16 is highlighted . As each portion of the text is converted to audio output the controller determines if that text portion should be highlighted, if it should be highlighted, it delays, then controls the highlighting of the text portion. In this example the word "email" is highlighted

first. The estimated accuracy of the synthesis of "email" into speech at time t1 falls below a threshold and the controller, in response, controls the highlighting of the word, after a delay, at time t2. The highlighting of "email" has been removed by time t7, when the phrase "a bientot" is synthesised. The estimated accuracy 5 of the synthesis of "bientot" into speech at time t7 falls below a threshold and the controller controls the highlighting of the word "bientot", after a delay, at time t8.

In the preceding examples illustrated in Figures 3 to 6, the same type of highlighting (underscoring) was used. Of course, the different examples 10 illustrated in the Figures could each use a different type of highlighting. Furthermore, the controller could simultaneously perform any combination of the different examples on the same body of text.

Although a particular example of the present invention has been described in the preceding paragraphs, it should be appreciated that modifications and variations to the examples given can be made without departing from the scope of the invention as claimed.

What is claimed is:

1. An electronic device comprising:

a user interface having

a display for displaying text and

5 speech synthesiser means including a loudspeaker, arranged to convert an input, dependent upon a text, to an audio output representative of a person reading the text; and

control means, for controlling the display and for providing an input to the speech synthesiser means, arranged to control the display of a text, to provide an input
10 corresponding to the displayed text to the speech synthesiser, and to highlight a portion or portions of the displayed text, wherein the highlighting of a text portion is delayed with respect to the audio output corresponding to the text portion.

15 2. A device as claimed in claim 1 wherein the control means synchronises the highlighting with the conversion of text to audio output.

3. A device as claimed in claim 1 or 2 wherein the control means varies the highlighting with the conversion of text to audio output.

20 4. A device as claimed in any preceding claim wherein the control means extends the highlighting through the displayed text with the conversion of text to audio output.

25 5. A device as claimed in claim 4 wherein the control means unselectively extends the highlighting through all of the text.

6. A device as claimed in claim 4 or 5, wherein the extending of highlighting through a text portion lags the conversion of the text portion to audio output by the delay.

7. A device as claimed in any one of claims 4 to 6, wherein the highlighting extends discontinuously by portions of text corresponding to a word or words.

8. A device as claimed in any preceding claim wherein the control means
5 highlights a text portion for a limited duration.

9. A device as claimed in claim 8 wherein the highlighting isolates a portion or portions of text from a body of text, the initiation of the isolation of text lagging the conversion of the text to audio output by the delay and the isolation being
10 maintained for the limited duration.

10 A device as claimed in any preceding claim wherein the dimension of the highlighted portion is variable.

11. A device as claimed in any preceding claim wherein the dimension of the highlighted portion is a constant number of words long.

12. A device as claimed wherein the display displays the full sentence of text being converted.

20 13. A device as claimed in any preceding claim wherein the display displays the previous predetermined plurality of words that have been converted.

14. A device as claimed in any preceding claim further comprising a dictionary,
25 wherein the control means highlights words not in the dictionary.

15. A device as claimed in claim 14 wherein the control means provides different highlighting for the words not in the dictionary.

16. A device as claimed in any preceding claim wherein the control means is arranged to identify proper nouns in the text, and highlight them.

17. A device as claimed in claim 16 wherein the control means provides different highlighting for the proper nouns in the text.
5

18. A device as claimed in any preceding claim wherein the synthesiser means provides signals to the control means to effect the highlighting of specific words.

10 19. A device as claimed in claim 18 wherein the control means provides different highlighting to the specific words.

15 20. A device as claimed in any preceding claim wherein highlighting may comprise any one of : contrast variation of the text and/or display background; colour variation of the text and/or display background; reformatting of the text including underscoring, bold font, italic font, capitalisation, changing font size or type; and enclosing text in geometric shapes.

20 21. A device as claimed in any preceding claim wherein the delay corresponds to a number of converted words or syllables.

22. A device as claimed in any preceding claim wherein the delay corresponds to a fixed time.

25 23. A device as claimed in any preceding claim wherein the delay is greater than 0.1 seconds.

24. A device as claimed in any preceding claim wherein the delay is less than 3 seconds.

25. A device as claimed in any preceding claim wherein the electronic device is a document reader or a hand-held communications device.

26. A hand-held radio communications device comprising:

5 a user interface having:

a display for displaying text and

speech synthesiser means including a loudspeaker, arranged to convert an input, dependent upon a text, to an audio output representative of a person reading the text; and control means for controlling the display and for providing 10 an input to the speech synthesiser means, arranged to control the display of a text, to provide an input corresponding to the displayed text to the speech synthesiser, and to highlight a portion or portions of the displayed text, wherein the highlighting of a text portion is delayed with respect to the audio output corresponding to the text portion.

15

27. A document reader comprising a user interface having:

a display for displaying text and

speech synthesiser means including a loudspeaker, arranged to convert 20 an input, dependent upon a text, to an audio output representative of a person reading the text; and control means for controlling the display and for providing an input to the speech synthesiser means, arranged to control the display of a text, to provide an input corresponding to the displayed text to the speech synthesiser, and to highlight a portion or portions of the displayed text, wherein 25 the highlighting of a text portion is delayed with respect to the audio output corresponding to the text portion.

28. A car having a device as claimed in claim 26 or a document reader as claimed in claim 27, wherein the display is mounted for use by the driver.

29. A device or circuitry substantially as hereinbefore described with reference to the accompanying figures and/or as shown in the figures.

30. A method for displaying text and providing speech synthesis of the text

5 comprising the steps of:

displaying the text;

converting a text portion to audio output;

determining that the text portion should be highlighted;

delaying; and then

10 highlighting the text portion.

31. An electronic device comprising:

a user interface having:

15 a display for displaying text and

speech synthesiser means including a loudspeaker, arranged to convert an input, dependent upon a text, to an audio output representative of a person reading the text; and

control means, for controlling the display and for providing an input to the speech synthesiser means, arranged to control the display of a text, to provide an input corresponding to the displayed text to the speech synthesiser, and to delay the display of a text portion with respect to the audio output corresponding to the text portion.

25 32. A method for displaying text and providing speech synthesis of the text comprising the steps of:

converting a text portion to audio output;

delaying; and then

displaying the text;

33. A device as claimed in claim 31 or a method as claimed in claim 32 where the delay is greater than 0.1 seconds.

Abstract

- An electronic device comprising: a user interface having a display for displaying
5 text and a speech synthesiser including a loudspeaker, arranged to convert an
input, dependent upon a text, to an audio output representative of a person
reading the text. The device may delay with respect to the audio output
corresponding to a portion of text the display of the portion of text. The device
may also or alternatively delay with respect to the audio output corresponding to
10 a portion of text the highlighting of the portion of text within a displayed text.

Document - Disclosure

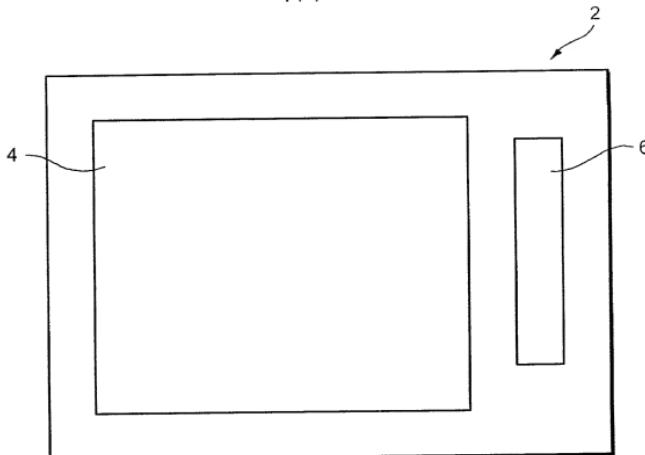


FIG. 1

DRAFTS - DRAWINGS

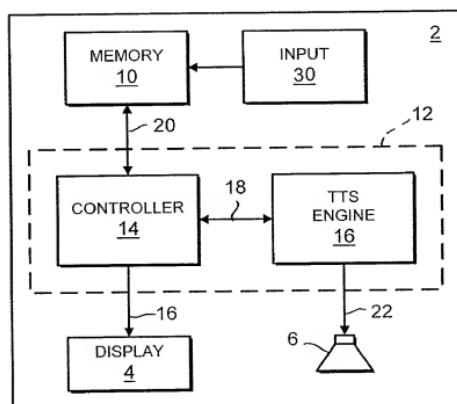


FIG. 2

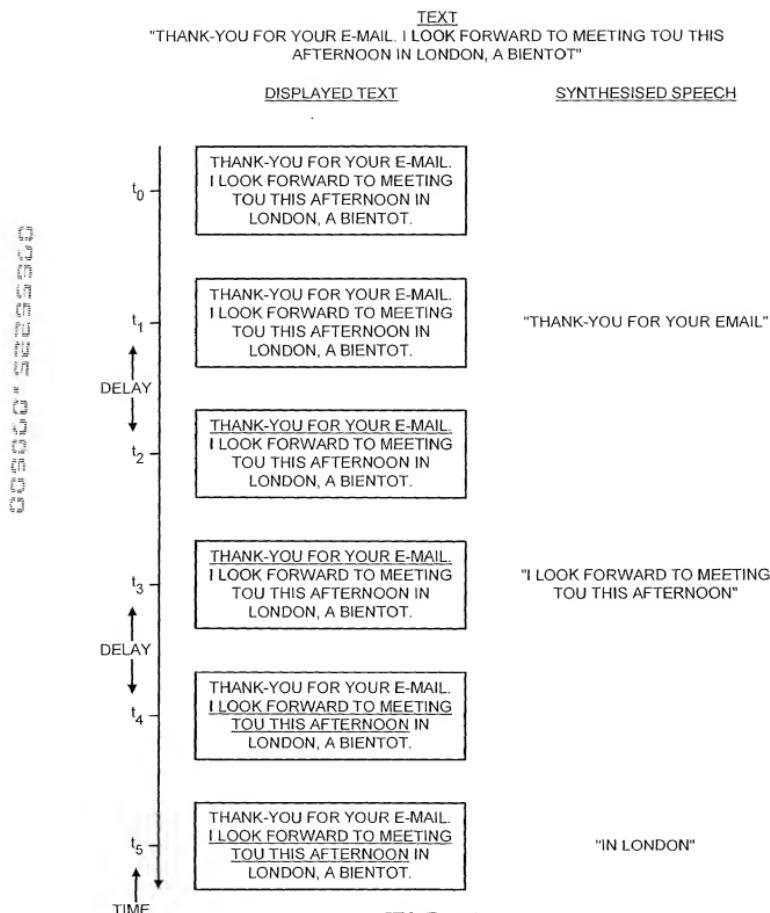


FIG. 3

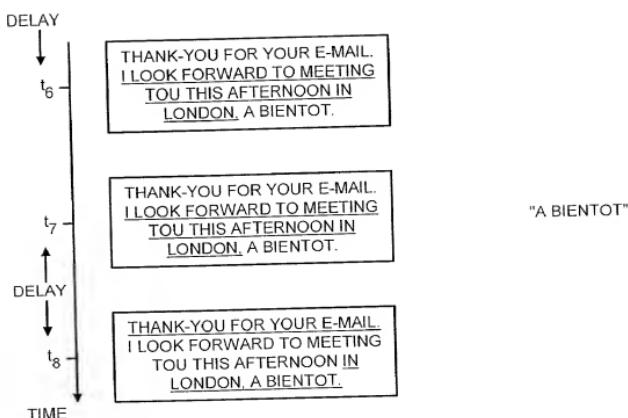


FIG. 3
(CONTINUED)

DISCUSSIONS • DOCUMENTS • IDEAS

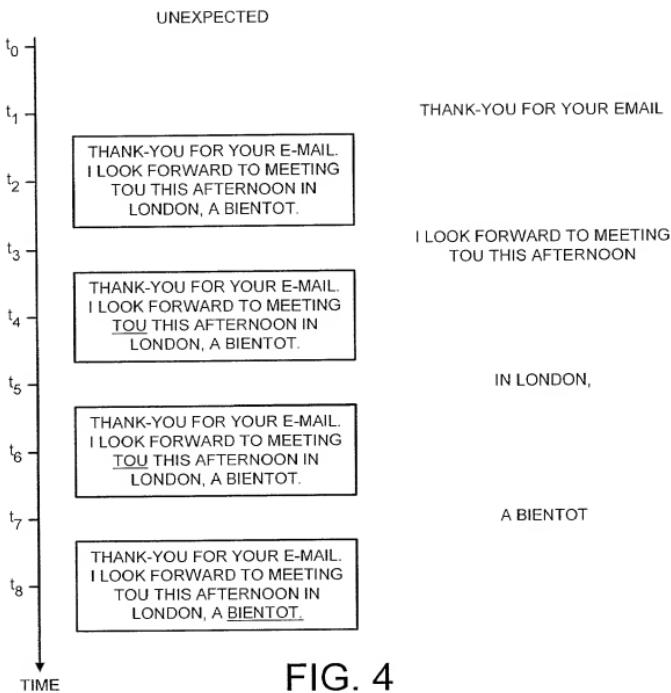
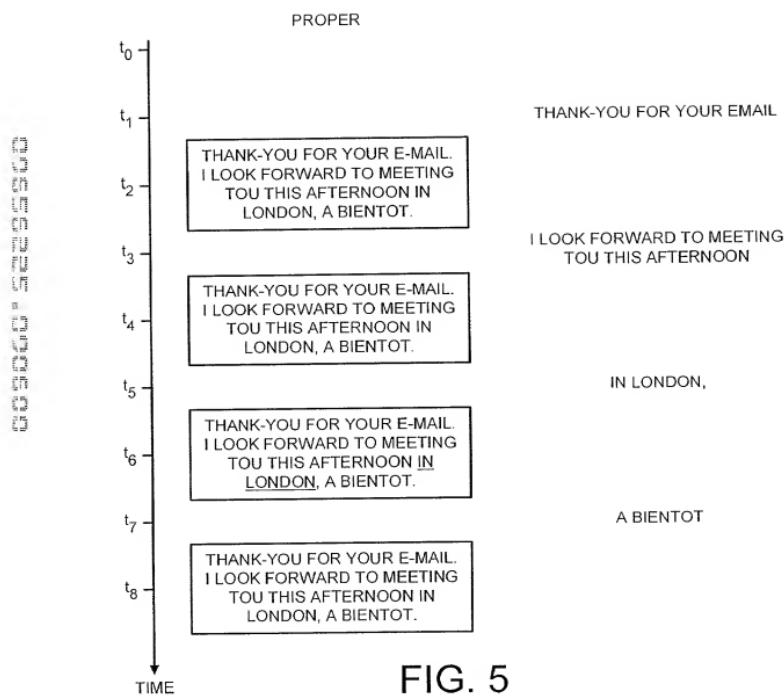


FIG. 4



MESSAGE 333333

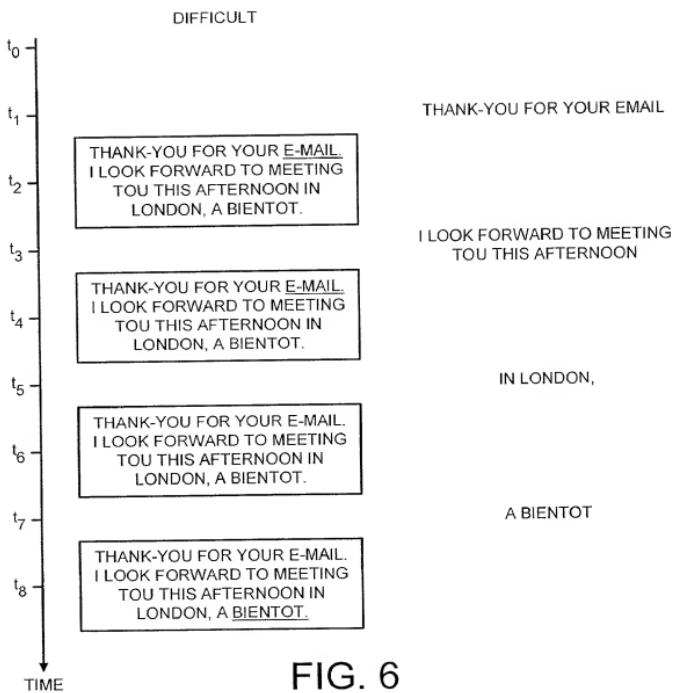


FIG. 6

The diagram illustrates a sequence of 16 words, each associated with a unique identifier (S_1 through S_{16}), a word size (1 or 2), a text group ID (1, 2, or 3), a selective highlight type (DIFFICULT, UNEXPECTED, or PROPER), and a time of audio output (t_1 through t_7). Four arrows point from numerical values (32, 34, 36, 38) to the words 'IN', 'LONDON,', 'A', and 'BIENTOT.' respectively, indicating their positions in the sequence.

	WORD SIZE	TEXT GROUP ID	SELECTIVE HIGHLIGHT TYPE	TIME OF AUDIO OUTPUT
THANK-YOU	S_1	1		t_1
FOR	S_2	1		t_1
YOUR	S_3	1		t_1
E-MAIL.	S_4	1	DIFFICULT	t_1
I	S_5	2		t_3
LOOK	S_6	2		t_3
FORWARD	S_7	2		t_3
TO	S_8	2		t_3
MEETING	S_9	2		t_3
TOU	S_{11}	2	UNEXPECTED	t_3
THIS	S_{12}	2		t_3
AFTERNOON	S_{12}	2		t_3
IN	S_{13}	3		t_5
LONDON,	S_{14}	3	PROPER	t_5
A	S_{15}	4		t_7
BIENTOT.	S_{16}	4	UNEXPECTED DIFFICULT	t_7

FIG. 7